ELSEVIER

Analytica Chimica Acta 433 (2001) 135–143

ANALYTICA CHIMICA ACTA

www.elsevier.com/locate/aca

Maximum likelihood estimation with nonlinear regression in polarographic and potentiometric studies

J.L.G.C. Pereira, A.A.C.C. Pais*, J.S. Redinha

Departamento de Química, Universidade de Coimbra, P-3049 Coimbra Codex, Portugal

Received 9 June 2000; received in revised form 15 November 2000; accepted 21 November 2000

Abstract

In this work we review some aspects of maximum likelihood nonlinear modeling in polarographic and potentiometric techniques. Different algorithms, namely the Levenberg–Marquardt and the "error-in-variables" methods in parametric and Monte-Carlo nonparametric estimation are used. Conclusions are drawn upon the influence of experimental errors and error correlation, introduced via statistical weighting, in the accuracy and precision of the estimated parameters. Several of the tested alternatives, including regression on the signal variable alone with a global error weighting function, are shown to provide adequate representation of the experimental data. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Maximum likelihood; Nonlinear regression; Polarographic studies

1. Introduction

The use of combined polarographic and potentiometric methods provides a powerful means to characterize complexometric systems and obtain data allowing for the determination of conditional formation constants [1]. This ultimate goal can only be achieved through a careful analysis of the successive steps in the estimation process so that accurate parameter values and respective uncertainties may be obtained.

The critical tasks on parameter estimation are related to model selection, and experimental error analysis [2,3]. These must follow optimally designed experimental procedures with data treatment in view. This work concerns experimental error analysis and problems related to statistical weighting of data and their effect on parameter estimation.

Obtaining experimental data suitable for a a posteriori error analysis is time consuming, especially when there is the need for "blank runs" to access experimental uncertainty. This assessment may also pose various problems when experiments are unique, e.g. in destructive essays, and/or simultaneous determination of errors in all variables is not possible due to instrumental interference.

Considering stable experimental conditions and reproducible instrumental response the above difficulties can be circumvented using a functional representation of the experimental error. This expression can then be employed in the statistical weighting of data obtained under similar conditions.

In the present work we discuss strategies to perform the above task and suggest alternatives to improve data modeling and parameter estimation. The structure of this article is as follows. Section 2 reviews the maximum likelihood formalism and discusses

* Corresponding author. Tel.: +351-239-852080; fax: +351-239-827703.
E-mail address: pais@qui.uc.pt (A.A.C.C. Pais).

0003-2670/01/$ – see front matter
PII: S0003-2670(00)01320-9

the influence of experimental error in parameter estimation, Section 3 describes the polarographic and potentiometric models used in case studies and the Monte-Carlo procedures for error simulation, Section 4 presents the most relevant results and Section 5 gathers some conclusions.

2. Maximum likelihood estimation

Polarographic and potentiometric curves may both be addressed as bivariate systems in which a perturbation is imposed, corresponding to the independent variable with value x_i, mean μ_{x_i} and random error ε_{x_i}, while the response is recorded as the dependent variable with value y_i, mean μ_{y_i} and random error ε_{y_i}.

In what follows we will assume that no significant systematic errors are present in the variables, being the main contribution to data dispersion additive random heterocedastic normal errors in the signal variable. This implies error free or, more loosely, abscissae affected by small homocedastic errors. We will also assume that a suitable and parsimonious model ($\eta(x_i, \theta_j)$) with p nonrandom parameters θ_j for the n data points (x_i, y_i) is available.

The deviation of the model in each data point is then given by

$$e_i = y_i - \eta_i, \quad \forall i \in \{1, \ldots, n\} \tag{1}$$

where η_i is the value predicted from the model for the dependent variable value y_i.

For a correct model no bias should occur which leads to the equality of the expected values for the model and experimental errors,

$$E(e_i) = E(\varepsilon_{y_i}) \tag{2}$$

In these conditions, each data ordinate follows a conditional normal distribution [2],

$$f(y_i|x_i) = \frac{1}{\sqrt{2\pi\sigma_{y_i}^2(1-\rho_{[x_i y_i]}^2)}} \times \exp\left[-\frac{\left(z_{y_i} - \rho_{[x_i y_i]} z_{x_i}\right)^2}{2\left(1-\rho_{[x_i y_i]}^2\right)}\right] \tag{3}$$

where

$$z_{x_i} = \frac{x_i - \mu_{x_i}}{\sigma_{x_i}}$$

and

$$z_{y_i} = \frac{y_i - \mu_{y_i}}{\sigma_{y_i}}$$

are the normalized random errors in the independent and dependent variables, respectively, and $\rho_{[x_i y_i]}$ is the experimental correlation between these variables. We note the conditional probability distribution function in Eq. (3) is used essentially to emphasize the existence of random error correlation between the independent and dependent variables.

Assuming that the normal error present in the independent variable propagates via the physically-based model to the corresponding predicted value (η_i), the model error of Eq. (1) is thus also characterized by a conditional normal distribution,

$$f(e_i|x_i) = f(y_i - \eta_i|x_i) = \frac{1}{\sqrt{2\pi\sigma_{e_i}^2\left(1-\rho_{[x_i e_i]}^2\right)}} \times \exp\left[-\frac{(z_{e_i} - \rho_{[x_i e_i]} z_{x_i})^2}{2\left(1-\rho_{[x_i\, e_i]}^2\right)}\right] \tag{4}$$

where

$$z_{e_i} = \frac{y_i - \eta_i}{\sigma_{e_i}}$$

is the normalized error of the model and $\rho_{[x_i e_i]}$ is the correlation between the independent variable and the model propagated error for the dependent variable.

When the errors in the independent variable are not significant in comparison to those present in the dependent variable,[1] $\rho_{[x_i e_i]} z_{x_i} \ll z_{e_i}$, Eq. (4) simplifies to

$$f(e_i|x_i) = \frac{1}{\sqrt{2\pi\sigma_{e_i}^2}} \exp\left(-\frac{z_{e_i}^2}{2}\right) \tag{5}$$

We note that for favorable situations in least-squares modeling, the distribution given by Eq. (5) is followed at all data points. If errors are uncorrelated for

[1] This situation is not infrequent when data are collected with adequate experimental/instrumental conditions.

different data points, the total conditional probability is given by the likelihood function ($L(\theta)$) [2],

$$L(\theta) = \prod_{i=1}^{n} f(e_i|x_i) \tag{6}$$

Best estimates of parameter vector θ can be found by maximizing the log-likelihood function ($l(\theta)$),

$$l(\theta) = -\frac{n}{2}\ln(2\pi) - \ln\left[\sum_{i=1}^{n}\sigma_{e_i}\right] - \frac{1}{2}\sum_{i=1}^{n}\left[\frac{(y_i-\eta_i)^2}{\sigma_{e_i}^2}\right] \tag{7}$$

which corresponds to minimize the weighted sum of squares term,

$$\text{SSqr} = \sum_{i=1}^{n}\left[\frac{(y_i-\eta_i)^2}{\sigma_{e_i}^2}\right] \tag{8}$$

However, variance can be obtained by general error propagation formulas,

$$\sigma_{e_i}^2 = \sigma_{y_i}^2 + \left(\frac{\partial\eta_i}{\partial x_i}\right)^2\sigma_{x_i}^2 - 2\left(\frac{\partial\eta_i}{\partial x_i}\right)\sigma_{y_i}\sigma_{x_i}\rho_{[x_i\,y_i]} \tag{9}$$

leading to an alternate form of Eq. (8).

$$\text{SSqr} = \sum_{i=1}^{n}\left[\frac{(y_i-\eta_i)^2}{a+b\delta_i^2+c\delta_i}\right] \tag{10}$$

where δ_i ($=\partial\eta_i/\partial x_i$) is the first derivative of the model and the coefficients are related with pure instrumental error ($a \approx \sigma_{y_i}^2, b \approx \sigma_{x_i}^2$) and error correlation ($c = -2\sigma_{x_i}\sigma_{y_i}\rho_{[x_i\,y_i]}$).

We note that the error propagation expression, Eq. (9), is more appropriate for linear models. Also, the correlation of variable errors is often neglected, or unitary correlation assumed [4].

The uncertainty in each observed data point ordinate can alternatively be ascribed to the sum of two independent terms, one related to the actual intrinsical instrumental random error and the other to error propagation from the independent variable. The latter is, in this approach, the sole responsible for error correlation between variables and for the systems studied here, the former can be viewed as approximately constant. The expression for regression weights may thus be rewritten simply as [5]

$$\frac{1}{w_i} = \sigma_{y_i}^2 = \hat{\sigma}_y^2 + \left(\frac{\partial\eta}{\partial x_i}\right)^2\hat{\sigma}_x^2 \tag{11}$$

The parameters $\hat{\sigma}_y$ and $\hat{\sigma}_x$ in the above equation can be determined by least-squares, if a sufficient number of blank runs have been conducted. Although $\hat{\sigma}_x$ should correspond, for the systems under study, to the approximately homocedastic dispersion in the independent variable, for practical applications a significant deviation from this average dispersion may be found. One of the reasons is that the random fluctuation in y is usually dominant, and only for high absolute values of the first derivative does the ordinate dispersion deviate from the intrinsical constant value.

The need for a first derivative of the model in Eq. (11) can be accommodated by standard techniques, e.g. Savitzky–Golay derivatives [6] or iteratively updated model derivatives [7].

In cases where both variables have significant uncertainties, the likelihood function is frequently [8] written as

$$L(\theta) = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma_{x_i}}\exp\left(-\frac{z_{x_i}^2}{2}\right) \times \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma_{y_i}}\exp\left(-\frac{z_{y_i}^2}{2}\right) \tag{12}$$

which would lead to an expression similar to Eq. (7) but that is often used without considering correlation between the errors in the dependent and independent variables (as in Eq. (9)). This suggests that the uncertainty in x is not physically propagated to y. The use of Eq. (12) implies [8]

$$\text{SSqr} = \sum_{i=1}^{n}\left(\frac{y_i-\eta(x_i;\theta)}{d_i}\right)^2 \tag{13}$$

where

$$d_i = \frac{1}{w_i} = \sigma_{y_i}^2 + \left(\frac{\partial\eta}{\partial x_i}\right)^2\sigma_{x_i}^2 \tag{14}$$

Minimization of SSqr with the above weights corresponds to the “effective variance method”, in which errors in both variables are treated by ordinary (weighted) least-squares. We stress that the above

expression is not equivalent to Eq. (11) once the total signal variance σ_{y_i} in Eq. (14) differs from the intrinsical counterpart $\hat{\sigma}_{y_i}$.

Our purpose in what follows will be to assess the best procedure to determine parameter values and dispersion from experimental data. The tested procedures will include different weighting and fitting algorithms as well as alternatives in the synthetic data generation for nonparametric validation of the least-squares approaches.

3. Procedure

3.1. Modeling

The polarographic signal of a differential pulse experiment with small amplitude steps [9] can be modeled by

$$i = 4\theta_1 \left[\frac{\exp((E-\theta_2)/\theta_3)}{[1+\exp((E-\theta_2)/\theta_3)]^2} \right] + \theta_4 + \theta_5 E \quad (15)$$

where θ_1 and θ_2 are the most relevant polarographic parameters, related to peak current (i_p) and peak potential (E_p), respectively, θ_3 is related to the redox reversibility coefficient ($\alpha = RT/n_e F\theta_3$) and n_e is the number of electrons involved. The remaining parameters (θ_4 and θ_5) are introduced to account for background current and are indispensable for obtaining accurate and precise estimates of the relevant parameters [5].

The potential in a potentiometric titration curve of a diluted strong acid (HX) with a strong base (MOH) in the presence of an interfering ion (M^+) can be described by the Nicholsky equation [10],

$$E_{H^+} = \theta_1 + \theta_2 \log[[H^+] + \theta_3[M^+]] \quad (16)$$

where $[M^+]$ represents the total interfering ion concentration and $[H^+]$ the hydrogen ion concentration,

$$[H^+] = \tfrac{1}{2}\left(\phi + \sqrt{\phi^2 + 4\theta_4}\right) \quad (17)$$

with

$$\phi = \frac{C_B(\theta_5 - v_B)}{V} \quad (18)$$

In the above, C_B stands for strong base analytical concentration, v_B for the respective added volume and V for the total volume of solution contained in the titration vessel.

Parameters consist in the conditional galvanic cell potential constant (θ_1), the response sensitivity (θ_2), the conditional potentiometric selectivity constant for the interfering ion (θ_3), the conditional ionic product of water (θ_4) and the equivalent volume (θ_5).

The parameter values used in the simulation of data curves and data dispersion, Eq. (11), are summarized in Table 1. The way in which they were obtained will be discussed in the next subsections.

3.2. Data acquisition

To assess the influence of experimental errors in parameter and parameter dispersion estimation we have carried out a simulation based on actual experimental data curves. These correspond to 16 potentiometric titration curves, with ca. 200 points for each curve, and simultaneous determination of potential and volume of added base for each point. In the case of the polarographic curves we used 19 blank runs for establishing the signal (current intensity values) while the imposed potential, abscissae, was independently determined through 10 replicates of the 125 potential values scanned for each run, with each value measured with a high precision micromultimeter. We note that the blank runs were obtained from a lead perchlorate electrolyte solution and thus differ from the general situation in which the ligand molecule or ion is also present. Experimental standard error deviations in the dependent variables (σ_{y_i}) were obtained directly

Table 1
Parameter values used for Monte-Carlo simulation in potentiometric (upper line) and polarographic (bottom line) cases

θ_1	θ_2	θ_3	θ_4	θ_5	$\hat{\sigma}_x$	$\hat{\sigma}_y$
419.832 mV	58.5704 mV per decade	3.46209×10^{-12} M^{-1}	1.90541×10^{-14} M^2	0.416081 m	0.000553 ml	0.58170 mV
0.325977 μA	−0.393654 V	1.35462×10^{-2} V	-6.66148×10^{-3} μA	-2.13259×10^{-2} μA V^{-1}	0.000479 V	0.00126 μA

from the above data. In the case of the independent variable in the polarographic experiments, dispersion was assessed on the basis of several micromultimeter readings of the applied potential values. The volume precision for the microburette was calculated from mass determination and density values corresponding to the experimental conditions, using adequately purified water. Potentiometric titrations of strong acid ($HClO_4$) with strong base (NaOH) were conducted with a 665 DOSIMATE (Metrohm), with minimum volume increments of 0.001 ml, recorded with a pHM 95 (Radiometer) potentiometer (±0.1 mV) in a thermostatized vessel and in the presence of an inert atmosphere.

The polarographic determinations were carried out in a 693 VA Processor/694 VA Stand (Metrohm) polarograph using a 50 μM Pb^{2+} solution with 0.2 M $NaClO_4$, slightly acidified. The imposed potential in the polarographic determinations was measured with a 195A Keithley digital multimeter (±10 μV).

3.3. Simulation

In this work we have generated synthetic data points to mimic experimental data. When the latter correspond to a set of calibration curves, the dispersion of the independent and signal variables may be assessed and expressed through an appropriate error structure to be used in similar experimental conditions when it is hard to get replicates. It is therefore of paramount importance to test the accuracy and precision of parameters recovered from a single data curve, with characterized dispersion in each point. In laboratory-based work, the data curve is replaced by experimental results, and the available dispersion comes from the error function determined in a corresponding calibration procedure [5]. Thus, parameters pertaining to Eq. (11) were determined on the basis of a linear unweighted least-squares to standard errors of experimental results.

The Monte-Carlo simulation we have used is based on experimental data that is previously subject to a fitting procedure (see Figs. 1 and 2). This procedure is intended to be completely dissociated from the subsequent tests, that comprise alternatives for the generation of synthetic data, weighting, and fitting algorithm. The parameters in Eqs. (15) and (16), describing the potentiometric titration and polarographic curves,

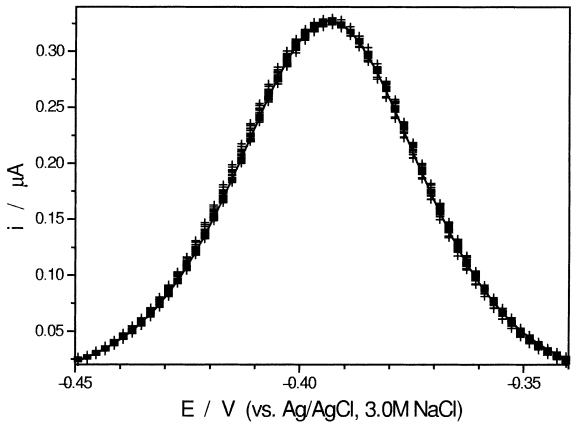

Fig. 1. Experimental differential pulse calibration points (19 replicates with average potential values) and model values (line), obtained with the Levenberg–Marquardt algorithm.

respectively, were obtained as follows: for each case, all available data (comprising the whole set of replicates) was fitted as only one data set, with equal weights. In the case of the polarographic data, average potential values were used as value of the independent variable in each point. This procedure yields parameter values for subsequent use in the simulation and the relative weight of each point is implicitly introduced through the experimental dispersion. Both the Levenberg–Marquardt [11] and "error-in-variables" [12] fitting algorithms were used. The first algorithm is directed to nonlinear least-squares and assumes that dispersion is located only in the dependent variable of the function subject to modeling. In contrast, the latter encompasses dispersion in both variables and may thus be used in nonlinear problems in which random errors present in the independent variable cannot be neglected. For simplicity, the values presented in Table 1 are the arithmetic mean of the results of these two algorithms.

Each Monte-Carlo simulation corresponds to a set of 2000 curves a number that is more than sufficient to guarantee adequate convergence of the nonparametric estimates. In the generation of these synthetic data sets (x_i, y_i), noncorrelated error was simulated using independent normal random number generators,

$$R_1 = \mu_{x_i} + N(0, \sigma_{x_i}^2) \tag{19}$$

$$R_2 = \eta(x_i, \theta_j) + N(0, \sigma_{y_i}^2) \tag{20}$$

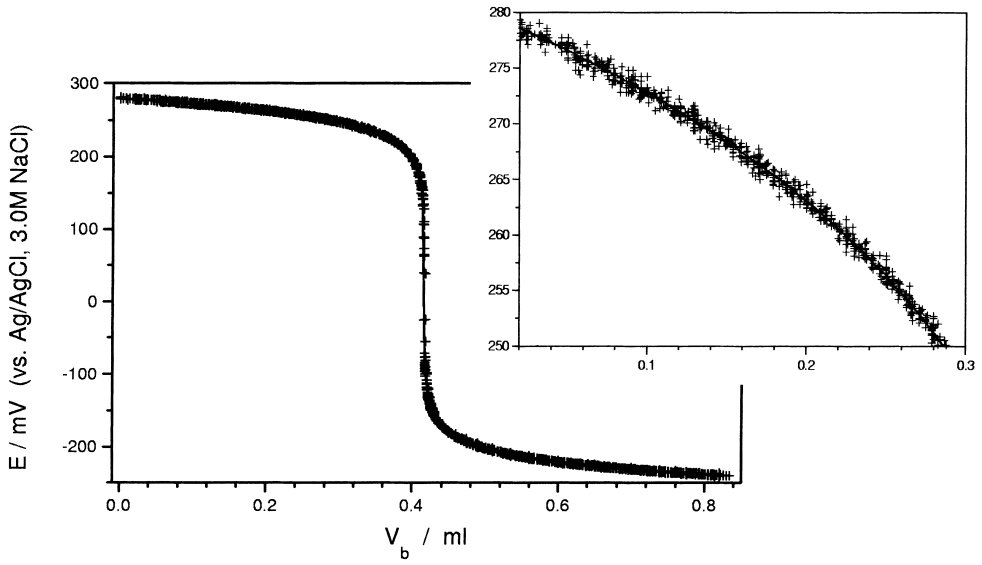

Fig. 2. Experimental potentiometric calibration points (16 acquisitions) and fitted values, using the Levenberg–Marquardt algorithm (line). The insert above highlights the typical data dispersion.

In the possibility of correlated error, also tested, the randomized value of the independent variable is used to generate the randomized dependent value,

$$R_2 = \eta(R_1, \theta_j) + N(0, \hat{\sigma}_y^2) \tag{21}$$

Besides using Eq. (11) as weighting factor, other functions were also tested for comparison. These correspond to (i) constant, unitary, weight ($w_i = 1$), (ii) constant relative error ($w_i = 1/y_i^2$), (iii) experimental y_i variance ($w_i = 1/\sigma_{y_i}^2$) and (iv) the effective variance factor of Eq. (14).

We note that weights (i) and (ii) above do not allow for direct consideration of experimental error propagation to parameters, and will be essentially used for accuracy assessment. Weight (iii) can only be applied in calibration procedures, and not in the actual experimental determinations.

Nonlinear regression on y_i is considered for all the weighting factors. In the case of the "error-in-variables" algorithm [12] only Eq. (11) and weight (iii) are presented. This is because the "effective variance method" encompasses errors in both variables in a ordinary (weighted) least-squares procedure. Explicit regression on both variable is conducted in this algorithm by minimizing the merit function

$$\text{SSqr} = \sum_{i=1}^{n} w_x (x_i - \hat{x}_i)^2 + w_{y_i} (y_i - \eta(\hat{x}_i))^2 \tag{22}$$

where $w_x = 1/\sigma_x^2$, corresponding directly to the inverse of the approximately homocedastic variance in the perturbation variable.

4. Results and discussion

The accuracy results for parameters recovered with Monte-Carlo simulation are presented in Table 2 while the corresponding precision values can be found in Table 3.

Firstly, we compare the "error-in-variables" results using Eq. (11) and the experimental variance for the signal variable. Both in the potentiometric and polarographic cases it is patent that the accuracy resulting from the use of both weighting methods is very similar. Also, it is obvious that the *xy* correlated generation of synthetic data, through Eq. (21), leads to better estimates of the parameters. Thus, in the following discussion we will concentrate on the correlated approach.

Table 2
Accuracy of parameters obtained by Monte-Carlo simulation, using the Levenberg–Marquardt and "error-in-variables" algorithms[a]

	Levenberg–Marquardt					"Error-in-variables"		
	Eq. (14)	$1/\sigma_{y_i}^2$	Eq. (11)	$1/y_i^2$	$1/\hat{\sigma}_y^2$	$1/\sigma_x^2, 1/\sigma_{y_i}^2$	Eq. (11)	
θ_1	−0.00028	−0.00028	−0.00028	−0.00027	−0.00028	−0.00034	−0.00034	Parametric estimate
	−0.00131	−0.02244	−0.00122	−0.03573	−0.09118	−0.11462	−0.08537	MC/ind.
	0.00803	−0.00056	−0.00015	0.00624	0.00736	−0.00022	−0.00112	MC/corr.
θ_2	0.00005	0.00005	0.00005	0.00005	0.00005	−0.00012	−0.00012	Parametric estimate
	−0.01055	−0.06674	−0.01368	−0.04612	−0.22423	−0.28595	−0.24421	MC/ind.
	0.02098	0.01056	0.01077	0.02368	0.01838	−0.00757	−0.01369	MC/corr.
θ_3	−0.00250	−0.00258	−0.00265	−0.00309	−0.00305	−0.01027	−0.01027	Parametric estimate
	−0.58918	−0.76350	−0.46810	3.31031	−6.95310	−0.40441	−0.94313	MC/ind.
	−0.47237	−0.33974	−0.27928	0.55336	0.32472	0.12952	0.12658	MC/corr.
θ_4	0.22181	0.22181	0.22178	0.22198	0.22179	0.21899	0.21899	Parametric estimate
	−0.07445	−0.92311	0.02257	−1.47437	−4.04343	−7.60489	−5.26055	MC/ind.
	0.14648	0.04925	0.17469	0.75899	0.50796	0.21915	0.14149	MC/corr.
θ_5	0.00002	0.00002	0.00002	0.00002	0.00002	−0.00002	−0.00002	Parametric estimate
	0.00045	−0.01166	0.00024	0.01540	−0.00411	−0.08652	−0.05761	MC/ind.
	0.00086	0.00053	0.00009	−0.00012	−0.00007	0.00032	0.00050	MC/corr.
θ_1	−0.03742	−0.01093	−0.01003	0.02400	−0.00893	−0.01058	−0.01189	Parametric estimate
	−0.04128	0.00137	−0.05780	13.49824	−0.05130	0.03772	−0.05780	MC/ind.
	−0.01019	−0.01679	−0.01749	13.42446	−0.01216	−0.00734	−0.01817	MC/corr.
θ_2	0.00211	−0.00184	−0.00189	0.33073	−0.01162	0.00702	0.00755	Parametric estimate
	−0.00969	−0.00962	−0.00964	0.99439	−0.00967	−0.00911	−0.00939	MC/ind.
	0.00047	0.00009	0.00042	1.00310	0.00072	0.00047	0.00047	MC/corr.
θ_3	−0.17149	−0.12995	−0.00340	1.44666	0.60958	−0.18602	−0.12243	Parametric estimate
	−0.00767	−0.05695	0.00216	0.46019	0.01605	0.02299	0.02041	MC/ind.
	0.03943	0.02604	0.04201	−0.00192	0.04644	0.04201	0.03389	MC/corr.

[a] Values are expressed in percentage, relative to those in Table 1. Various weighting factors are used, as indicated on the top of each value column. The upper part of the table corresponds to the potentiometric and the lower part to the polarographic data. Key — MC/ind.: Monte-Carlo estimates with independent generation of synthetic x and y data, Eqs. (19) and (20); MC/corr.: Monte-Carlo estimates using Eqs. (19) and (21).

The use of the simpler Levenberg–Marquardt instead of the "error-in-variables" approach slightly increases the accuracy of the results obtained from the weighting function (11). We note, however, that both are essentially equivalent and that the original generating parameters were defined by the arithmetic mean of results from those algorithms, obtained from the complete set of experimental data. The effective variance weighting leads to parameters not so accurate as those stemming from Eq. (11) in the case of the potentiometric titrations, although slightly increasing the accuracy in the polarographic curves (for which only average values in each abscissa were used in the determination of the original, i.e. generating parameters of Table 1). The consideration of equal weights in all points yields for the two systems considered an accuracy similar to that obtained with the "effective variance method" or Eq. (11). This conclusion does not apply to weighting based on constant relative error ($w_i = 1/y_i^2$), that in some parameters presents considerable deviations from the values of Table 1. Probably, the accuracy obtained with constant weighting is due to the fact that most experimental and, consequently, simulated points are located in regions were the first derivative of the model function has a small value.

In what concerns the respective parameter dispersion, the situation is different. Constant weighting predicts parameter dispersions very different from every other estimate. This leads to the obvious result that even if this approach may be considered in accuracy terms, it must be discarded on precision

Table 3
Precision of parameters obtained by Monte-Carlo simulation, using the Levenberg–Marquardt and "error-in-variables" algorithms[a]

	Levenberg–Marquardt					"Error-in-variables"		
	Eq. (14)	$1/\sigma_{y_i}^2$	Eq. (11)	$1/y_i^2$	$1/\hat{\sigma}_y^2$	$1/\sigma_x^2, 1/\sigma_{y_i}^2$	Eq. (11)	
θ_1 (mV^{-1})	0.617	0.570	0.562	–	0.363	0.570	0.562	Parametric estimate
	0.627	0.544	0.564	2.786	2.431	0.543	0.540	MC/ind.
	0.608	0.558	0.561	0.384	0.368	0.562	0.549	MC/corr.
θ_2 (decade mV^{-1})	0.232	0.212	0.209	–	0.127	0.212	0.209	Parametric estimate
	0.235	0.224	0.232	1.042	0.926	0.210	0.227	MC/ind.
	0.250	0.217	0.219	0.134	0.129	0.208	0.207	MC/corr.
$\theta_3 \times 10^{12}$ (M)	0.589	0.551	0.554	–	0.352	0.551	0.539	Parametric estimate
	0.598	0.617	0.622	3.531	2.456	0.951	0.843	MC/ind.
	0.533	0.546	0.549	0.358	0.338	0.526	0.528	MC/corr.
$\theta_4 \times 10^{15}$ (M^{-2})	1.317	1.201	1.180	–	0.710	1.156	1.090	Parametric estimate
	1.339	1.369	1.371	5.991	4.913	3.389	2.548	MC/ind.
	1.181	0.987	1.073	0.749	0.726	1.345	1.365	MC/corr.
$\theta_5 \times 10^3$ (ml^{-1})	0.252	0.210	0.196	–	0.021	0.208	0.197	Parametric estimate
	0.261	0.272	0.268	0.589	0.604	0.270	0.270	MC/ind.
	0.235	0.212	0.213	0.025	0.024	0.213	0.211	MC/corr.
$\theta_1 \times 10^2$ (μA^{-1})	0.084	0.071	0.069	–	0.040	0.068	0.069	Parametric estimate
	0.073	0.072	0.073	1.883	0.120	0.072	0.071	MC/ind.
	0.068	0.071	0.065	1.903	0.087	0.071	0.068	MC/corr.
$\theta_2 \times 10^3$ (mV^{-1})	0.100	0.075	0.076	–	0.028	0.081	0.080	Parametric estimate
	0.096	0.078	0.076	2.553	0.121	0.077	0.073	MC/ind.
	0.081	0.078	0.075	2.573	0.089	0.077	0.073	MC/corr.
$\theta_3 \times 10^4$ (V^{-1})	0.494	0.418	0.420	–	0.231	0.423	0.421	Parametric estimate
	0.419	0.406	0.412	1.796	1.704	0.410	0.420	MC/ind.
	0.417	0.409	0.410	1.795	1.702	0.414	0.416	MC/corr.

[a] Various weighting factors are used, as indicated. The upper data corresponds to potentiometric and the lower part to polarographic data. Key as in Table 2.

estimate grounds. Further, the use of Eq. (11) and the experimental-based weights yield very similar results, whilst the effective variance weighting leads to higher dispersion estimates. These conclusions reflect the fact that error correlation between the perturbation and signal variables is not accounted for in Eq. (14).

The parametric estimate of dispersion, as pointed out before [5], is an acceptable choice as alternative for the Monte-Carlo correlated simulation. Parameter errors using this simple approach are only slightly higher than those obtained through the more computationally expensive simulation methods.

The use of the model derivative in the weighting functions (11) and (14) also deserves some comment. A trial-and-error procedure has shown that although the tested alternatives, Savitzky–Golay [6] and iteratively-reweighting [7], lead essentially to the same solutions, they behave differently in terms of convergence. An optimized procedure consists in using the Savitzky–Golay first derivative in the initial iteration of the least-squares procedure, thus avoiding problems caused by incorrect initial guess values of parameters. The subsequent steps are conducted with numerical derivatives iteratively updated. This ensures a faster convergence rate and fewer divergence problems than maintaining the Savitzky–Golay values in the whole process.

5. Conclusions

The first conclusion that can be drawn, and this comes in the sequence of our previous work [5], is that the fitting in both variables is not relevant for systems such as those describe in here, for the obvious reason

that the order of magnitude of the error in the perturbation variable is considerably lower than that corresponding to the signal variable. Another conclusion is that several approaches in the weighting function can be used to accurately estimate the relevant parameters, namely, Eqs. (11) and (14) or even the constant weighting function, although the latter must be viewed with extreme care. The precision estimates are much more dependent on the weighting procedure. In our view, Eq. (11) should be employed for two reasons: first it is easy to transpose to similar experimental condition and, secondly, it yields results similar to weighting based on experimental data. Also, it conveniently and accurately describes the heterocedastic nature of the signal variable for both the potentiometric titration and polarographic methods. For these methods, extension to heterocedastic independent variables was not deemed necessary. The effective variance method enlarges parameter error estimates, but is still a valid approach especially for conservative parameter error estimates. It must, however, be used in conjunction with an appropriate error description for conditions in which it is difficult to determine experimental errors. If one is seeking the simplest method, it must be stressed that parametric estimates are, for both the potentiometric and polarographic techniques, very reliable in what concerns precision determinations.

Acknowledgements

The experimental part of this work was conducted in ÆMITEQ — Associação para a Inovação Tecnolgica e Qualidade, Coimbra, Portugal.

References

[1] I. Cukrowski, Anal. Chim. Acta 336 (1996) 23.
[2] J.V. Beck, K.J. Arnold, Parameter Estimation in Engineering and Science, Wiley, New York, 1977.
[3] T.P. Ryan, Modern Regression Methods, Wiley, New York, 1997.
[4] D.P. Shoemaker, C.W. Garland, Experiments in Physical Chemistry, McGraw-Hill, New York, 1962.
[5] A.A.C.C. Pais, J.L.G.C. Pereira, J.S. Redinha, Comput. Chem. 24 (2000) 533.
[6] A. Savitzky, M.J.E. Golay, Anal. Chem. 36 (1964) 1627.
[7] G.R. Philips, E.M. Eyring, Anal. Chem. 55 (1983) 1134.
[8] J. Orear, Am. J. Phys. 50 (1982) 912.
[9] J. Churacek, P.J. Cox, M.R. Masson, Advanced Instrumental Methods of Chemical Analysis, Ellis, New York, 1993.
[10] K. Camman, Working with Ion-Selective Electrodes, Springer, Berlin, 1979.
[11] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in Fortran: The Art of Scientific Computing, 2nd Edition, Cambridge University Press, Cambridge, 1992.
[12] P.M. Reilly, H.V. Reilly, S.E. Keeler, Appl. Statist. 42 (1993) 693.